

Validation of an observation instrument for assessing english language teachers' performance

Validación de un formato de observación para evaluar el desempeño docente en inglés

Carmen Leonor Ojeda Castillo, Arturo Llanes Castillo*

Abstract

Teacher performance evaluation plays a critical role in promoting the continuous improvement of educational practice. However, some evaluation instruments used in educational institutions fail to address the specific features of English language teaching, thereby limiting the provision of targeted feedback for teachers. The objective of this study was to examine the functionality of an observation instrument designed to assess teacher performance in English language instruction and to compare it with the institutional evaluation tool previously in use. A descriptive, cross-sectional, and comparative study was conducted at a private bilingual elementary school in Tampico, Tamaulipas, Mexico. The participants were 24 English teachers from first through sixth grade during the 2006–2007 academic year. An observation instrument based on Brown's framework was implemented through evaluations conducted by both internal and external observers. Subsequently, a questionnaire was administered to assess the instrument's usefulness during the feedback process. The findings demonstrated a high level of agreement between the two evaluations and revealed that 96% of the teachers obtained scores at or above the average benchmark. Furthermore, most participants reported that the instrument facilitated the identification of strengths and areas for improvement, thereby enhancing communication between observers and teachers. It is concluded that this instrument represents a valuable tool for strengthening formative assessment practices and supporting teachers' professional development.

Keywords: teacher performance evaluation; classroom observation; English language teaching; teacher professional development; formative feedback.

Resumen

La observación de aula constituye una estrategia fundamental para evaluar el desempeño docente y promover la mejora continua de la práctica educativa. No obstante, los instrumentos de evaluación utilizados en algunas instituciones no contemplan las particularidades de la enseñanza del idioma inglés, limitando la retroalimentación específica para los docentes. El objetivo de este estudio fue analizar la funcionalidad de un formato de observación diseñado para evaluar el desempeño docente en la enseñanza del inglés y compararlo con el instrumento institucional utilizado previamente. Se realizó un estudio descriptivo, transversal y comparativo en una institución privada de educación primaria bilingüe de Tampico, Tamaulipas. Participaron 24 docentes de inglés de primero a sexto grado durante el ciclo escolar 2006–2007. Se aplicó un formato de observación basado en la propuesta de Brown mediante evaluaciones realizadas por observadores internos y externos. Posteriormente, se administró un cuestionario para valorar la utilidad del instrumento durante la retroalimentación. Los resultados mostraron alta consistencia entre ambas evaluaciones y evidenciaron que el 96 % de las docentes obtuvo calificaciones iguales o superiores al promedio. Además, la mayoría consideró que el formato facilitó la identificación de fortalezas y áreas de mejora, favoreciendo la comunicación entre observador y docente. Se concluye que el instrumento constituye una herramienta útil para fortalecer la evaluación formativa y el desarrollo profesional docente.

Palabras clave: evaluación del desempeño docente; observación en el aula; enseñanza del inglés; desarrollo profesional docente; retroalimentación formativa.

Correspondencia: allanes@docentes.uat.edu.mx

Fecha de recepción: 09/febrero/2026 | **Fecha de aceptación:** 17/julio/2026 | **Fecha de publicación:** 04/agosto/2026

*Instituto Anglo Mexicano Primaria. Tampico, Tamaulipas, México



INTRODUCCIÓN

English language teaching has become a strategic component of contemporary educational systems due to its relevance to academic mobility, employability, and participation in increasingly globalized environments. Consequently, educational institutions face the ongoing challenge of ensuring high-quality instructional processes that effectively address students' language learning needs while fostering the development of communicative competencies aligned with the demands of an increasingly interconnected international context (Kobayashi & Kobayashi, 2024; Lewis & Wagner, 2023; Donaghue, 2026; Thida & Yang, 2026; British Council, 2025).

Within this context, teacher performance is recognized as one of the factors exerting the greatest influence on student learning outcomes. As a result, its evaluation and continuous improvement have become central concerns in contemporary educational policies and practices. Recent research indicates that systematic classroom observation provides direct evidence of pedagogical practice, enables the identification of strengths, areas for improvement, and professional development needs, and generates valuable information for teacher feedback. Furthermore, these processes contribute to enhancing educational quality through continuous improvement strategies (Padgett, 2023; Reños & Pontillas, 2024; Acar, Akgün Özpolat, & Çomoğlu, 2023; Amemasor et al., 2025).

Traditionally, teacher observation was primarily associated with supervision, accountability, and compliance functions. Over recent decades, however, a significant conceptual shift has occurred, moving away from inspection-oriented models toward approaches focused on professional development, critical reflection, and the collaborative construction of pedagogical knowledge. Classroom observation is now widely regarded as a formative tool that facilitates a deeper understanding of the complexities of teaching and learning through the systematic analysis of classroom interactions. In doing so, it supports the development of institutional cultures grounded in continuous improvement, ongoing professional learning, and evidence-informed educational practice (UCLA Teaching and Learning Center, 2025; Donaghue, 2026; British Council, 2025; Acar et al., 2023).

Empirical evidence further suggests that observation processes are most effective when integrated into structured cycles that include pre-observation planning, direct observation, evidence analysis, and reflective feedback. Such an approach moves beyond merely judging teacher performance and instead emphasizes understanding pedagogical practice and creating opportunities for professional growth. Post-observation dialogue, in particular, encourages critical reflection on teaching practices, evidence-based decision-making, and the implementation of actions aimed at continuous instructional improvement and the enhancement of educational quality (Donaghue, 2026; Padgett, 2023; Inwang, 2024; Tapia & Evison, 2024; Yaşar & Gürbüz, 2023).

In the specific field of English language teaching, classroom observation assumes particular significance because of the inherent complexity of foreign language learning. Effective instruction depends on multiple factors, including the selection of appropriate methodological strategies, the promotion of meaningful communicative interaction, the effective use of instructional resources, the management of corrective feedback, and the creation of learning environments that encourage active participation and the progressive development of students' linguistic competencies. Consequently, instruments designed to evaluate teacher performance in this area should incorporate specialized criteria capable of assessing dimensions unique to language teaching, rather than relying exclusively on general indicators that may be insufficient to capture the complexity and distinctive characteristics of professional practice in this discipline (Brown, 2014; Donaghue, 2026; Acar et al., 2023; Alghamdi, Al-Nofaie, & Alkhamash, 2023; Thida & Yang, 2026).

The quality of observation instruments represents a critical determinant of the validity and usefulness of evaluation processes. Well-designed instruments not only facilitate the systematic collection of relevant information regarding teacher performance but also enhance the objectivity, consistency, and relevance of classroom observations. In addition, they establish shared standards of judgment, strengthen communication between observers and teachers, and provide evidence to support feedback processes directed toward professional growth. Previous

studies have highlighted that the clarity of indicators, the specificity of evaluation criteria, and the alignment between observed elements and instructional objectives are essential factors for increasing the effectiveness, reliability, and practical value of teacher evaluation systems (Padgett, 2023; UCLA Teaching and Learning Center, 2025; Acar et al., 2023; British Council, 2025).

Moreover, recent literature emphasizes that classroom observation and teacher evaluation yield greater benefits when conducted within environments of professional trust and when perceived as opportunities for learning rather than mechanisms of sanction or control. Under such conditions, observation-based feedback promotes self-reflection, strengthens teachers' capacity for pedagogical analysis, encourages the development of more conscious, critical, and evidence-informed teaching practices, and transforms observation into a process of professional support oriented toward growth. These outcomes contribute to the development of teaching competencies, the strengthening of professional identity, and the establishment of institutional cultures centered on the continuous improvement of instruction (Donaghue, 2026; Reños & Pontillas, 2024; Thida & Yang, 2026; Acar et al., 2023).

Despite the conceptual and methodological advances achieved in the field of teacher evaluation, many educational institutions continue to employ observation instruments designed for general educational settings that do not adequately address

the specific requirements of English language teaching. This limitation may reduce the ability of such instruments to provide meaningful information on key pedagogical aspects related to foreign language learning, diminish their sensitivity in identifying discipline-specific strengths and areas for improvement, and restrict their potential to support effective professional development, pedagogical innovation, and evidence-based institutional decision-making (Brown, 2014; Padgett, 2023; Alghamdi et al., 2023; Acar et al., 2023; Donaghue, 2026).

Accordingly, it is important to examine the functionality and relevance of observation and evaluation instruments used in English language teaching contexts in order to determine their capacity to assess teacher performance accurately and contribute to educational quality improvement. The present study therefore analyzes an observation and evaluation instrument implemented in this disciplinary field, considering its structure, assessment criteria, and usefulness as a formative feedback tool. The study is based on the premise that an instrument grounded in sound pedagogical, linguistic, and assessment principles can serve as a strategic resource for strengthening teacher performance, promoting reflective teaching practices, supporting the continuous improvement of English language teaching and learning processes, and informing evidence-based institutional decision-making (Donaghue, 2026; Reños & Pontillas, 2024; Padgett, 2023; Thida & Yang, 2026; Amemasor et al., 2025).

METHODS, TECHNIQUES AND INSTRUMENTS

This study adopted a quantitative approach with a descriptive and comparative scope aimed at examining the functionality of an observation and evaluation instrument specifically designed for English language teaching. A cross-sectional design was employed, as data were collected during a defined period of the academic year, allowing for the description of the observed phenomenon and the comparison of the results obtained through the application of the teacher evaluation instrument within a specific educational context.

The study was conducted at the Anglo Mexicano Institute in Tampico, Tamaulipas, Mexico. The study population consisted of 24 teachers responsible for English language instruction from first through sixth grade in elementary education. During the research process, a specialized observation instrument for English language teaching was implemented to assess its usefulness in evaluating teacher performance and its contribution to the improvement of educational practice.

The primary data collection technique was direct and systematic classroom observation, considered an appropriate procedure for obtaining information about teaching practices in authentic instructional settings. The observations focused on aspects related to lesson preparation, content presentation, instructional methods and classroom implementation, teachers' personal characteristics,

and teacher-student interactions throughout the instructional process.

In addition to classroom observation, survey methodology was employed through the administration of a structured questionnaire directed to both the observed teachers and the participating observers. The purpose of this procedure was to gather information regarding the usefulness, clarity, relevance, and functionality of the observation and evaluation instrument used in the study, as well as participants' perceptions of its contribution to improving English language instruction.

Furthermore, pedagogical feedback was incorporated as a follow-up strategy after each observation, enabling teachers to receive specific information about their performance and reflect on the strengths and areas for improvement identified during the evaluation process.

Two primary instruments were used for data collection.

The first instrument was the Teacher Observation Form A: Observing Other Teachers, developed by Brown (1994) and adapted for evaluating English language teachers within the institutional context. The instrument consisted of 41 indicators organized into five dimensions:

- Lesson preparation
- Content presentation

- Instructional implementation and teaching methods
- Teacher personal characteristics
- Teacher-student interaction
-

Each indicator was rated using a four-point scale, where 4 represented excellent, 3 above average, 2 average, and 1 unsatisfactory. An additional Not Applicable option was included for aspects that could not be observed during the evaluated lesson. The instrument provided a comprehensive assessment of teacher performance based on criteria specifically related to English language teaching.

Following each observation and feedback session, a survey composed of ten dichotomous (Yes/No) questions was administered to assess teachers' perceptions of the instrument. The items explored aspects related to the instrument's usefulness, clarity of evaluation criteria, ability to identify strengths and areas for improvement, promotion of reflective dialogue, and appropriateness for evaluating English language teaching. The information obtained made it possible to assess the instrument's acceptance and functionality from the perspective of its direct users.

The research procedure began with the selection of participating teachers. Classroom observations were subsequently conducted using the specialized English language teaching observation instrument. Observations were carried out by both an internal observer and an external observer in order to compare results and enhance the objectivity of the

evaluation process. After each observation, an individual feedback session was held with the participating teacher, followed by the administration of the evaluation survey to determine the instrument's usefulness and benefits within the context of English language teaching.

Data obtained through the teacher observation and evaluation instrument and the participant survey were analyzed using descriptive statistical procedures. Frequencies, percentages, and mean scores were calculated to describe teacher performance across the evaluated dimensions and to determine teachers' perceptions of the instrument's usefulness. In addition, the results obtained by the internal and external observers were compared to assess the consistency of their evaluations. Finally, the data were organized and presented through tables and figures, facilitating the interpretation of findings and the formulation of conclusions regarding the functionality of the observation and evaluation instrument as a tool for improving English language teaching.

RESULTS AND DISCUSSION

The results obtained through the implementation of the teacher observation and evaluation instrument indicated that most participants demonstrated performance levels ranging from average to above average across the evaluated dimensions. Scores assigned by the external evaluator ranged from 1.92 to 3.29, whereas those assigned by the internal evaluator ranged from 2.04 to 3.39. Only minimal differences were observed between the two sets of

ratings, suggesting a satisfactory degree of consistency in the application of the instrument and a relatively homogeneous assessment of teacher performance by both observers.

In the lesson preparation dimension, which addressed classroom organization, instructional planning, and the establishment of learning objectives, teachers obtained average or above-average ratings. These findings suggest that lessons were adequately structured and that instructional planning was aligned with the intended educational goals. Similarly, results for the content presentation dimension reflected satisfactory performance in areas such as clarity of explanations, logical sequencing of content, comprehensibility of instructions, and the adaptation of instructional materials to students' proficiency levels, all of which are considered essential elements for effective English language learning.

Regarding the instructional implementation and teaching methods dimension, 96% of the teachers achieved above-average scores. Notable strengths included the use of varied learning activities, reinforcement of previously taught content, effective utilization of teaching resources, and the application of strategies for correcting language errors. These findings suggest that the instructional practices observed encouraged student participation and supported the development of communicative skills in the classroom, remaining consistent with contemporary principles of foreign language teaching.

With respect to teachers' personal characteristics, favorable results were found among all participants. Teachers obtained above-average ratings in aspects such as voice clarity and audibility, initiative, creativity, and appropriate use of English in terms of pronunciation, intonation, and fluency. These results indicate that the teachers' communicative and personal competencies represented a significant strength within the instructional process and contributed to creating learning environments conducive to student engagement and achievement.

Similarly, the teacher-student interaction dimension yielded highly positive results. All teachers received above-average ratings, demonstrating their ability to promote student participation, maintain classroom management, address individual learner needs, and create learning environments in which students remained attentive and actively engaged. The quality of teacher-student interaction is particularly important in English language instruction because it facilitates meaningful language use and the development of communicative competencies in authentic learning contexts.

The overall analysis showed that 96% of the teachers achieved above-average performance levels in the comprehensive evaluation. Only one teacher received a score below 2.0 in the external evaluator's assessment. This finding indicates that, overall, the participating teachers demonstrated satisfactory professional performance and that the observation instrument was capable of identifying specific

differences across varying levels of instructional effectiveness.

The results of the post-observation survey revealed a highly favorable perception of the observation and evaluation instrument. Twenty teachers responded positively to all ten survey items, while only four participants expressed disagreement with one or two specific statements. Most respondents indicated that the instrument facilitated the identification of strengths and areas for improvement, contributed to professional development, enhanced communication with the observer, and promoted reflective examination of teaching practice. Participants also emphasized that the instrument provided a more specific and meaningful evaluation of English language teaching than the institutional instrument previously used.

The findings demonstrate that implementing an observation instrument specifically designed for English language teaching enabled the systematic identification of strengths and areas for improvement in teacher performance. Performance levels were predominantly above average across the dimensions of lesson preparation, content presentation, instructional implementation, personal characteristics, and teacher-student interaction. These results are consistent with the findings of Reños and Pontillas (2024), who reported a significant relationship between classroom observation processes and teacher professional development, emphasizing that structured observation supports pedagogical improvement and

strengthens opportunities for continuous professional learning. Similarly, Padgett (2023) argues that observation instruments used for formative purposes constitute effective mechanisms for promoting improvements in educational practice by generating systematic processes of feedback and professional reflection.

The minimal differences observed between the assessments conducted by the internal and external observers suggest an acceptable level of consistency in the application of the instrument, thereby strengthening the reliability of the findings. This result aligns with the arguments presented by Donaghue (2026), who asserts that observation processes grounded in clearly defined criteria promote more objective interpretations of teacher performance and facilitate professional dialogue based on observable evidence rather than subjective impressions. Likewise, Acar, Akgün Özpolat, and Çomoğlu (2023) emphasize that structured observation instruments tailored to specific instructional contexts support consistent performance evaluations and enhance teachers' confidence in the credibility of observation processes.

The strong performance observed in the dimensions related to instructional implementation and teacher-student interaction reflects the presence of pedagogical practices that encourage active student engagement and foster effective learning environments. These findings are consistent with those reported by Thida and Yang (2026), who found

that classroom observation enables the identification of student-centered instructional practices characterized by communicative activities, responsiveness to individual learner needs, and active participation, all of which are fundamental components of effective English as a second language instruction. Similarly, Alghamdi, Al-Nofaie, and Alkhamash (2023) note that collaborative and formative observation models facilitate the identification of successful classroom interaction practices and communication-oriented methodologies, which are considered essential in foreign language education.

Survey responses further indicated a highly favorable perception of the observation instrument, particularly regarding the clarity of evaluation criteria, the identification of strengths and weaknesses, and the promotion of reflective dialogue following classroom observations. These results are consistent with Donaghue's (2026) review, which identifies post-observation feedback as one of the most influential factors in teacher professional development because it promotes critical reflection, professional learning, and the construction of new perspectives on educational practice. Furthermore, studies by Inwang (2024), Tapia and Evison (2024), and Yaşar and Gürbüz (2023), as cited by Donaghue (2026), suggest that structured feedback processes enhance professional confidence, increase commitment to continuous improvement, and contribute to the development of reflective teaching identities.

Another important finding concerned the perceived usefulness of the instrument for evaluating dimensions specific to English language teaching, a characteristic that contrasted with the limitations attributed to the previously used institutional format. Recent literature supports the need for specialized observation instruments tailored to particular teaching contexts. Acar et al. (2023) reported that instruments specifically designed for English language teachers provide more relevant information for pedagogical decision-making and support more meaningful observation processes than those based exclusively on general performance criteria. In addition, the British Council (2025) emphasizes that contemporary observation systems should consider the specific characteristics of both the educational context and the subject area in order to facilitate authentic assessments of professional performance and promote effective teacher development.

The findings also reinforce the view that classroom observation should be understood primarily as a professional development strategy rather than as an administrative control mechanism. This perspective is consistent with the framework proposed by the UCLA Teaching and Learning Center (2025), which argues that formative observation models generate greater benefits when they are perceived as opportunities for learning and collaborative reflection, fostering relationships of trust and cooperation between observers and teachers. Likewise, the systematic review conducted by Amemasor et al. (2025) concludes that the most effective professional development programs are

those that incorporate continuous support, contextualized feedback, and collaborative reflection, all of which were integral components of the procedure employed in the present study.

Taken together, the findings indicate that the observation and evaluation instrument analyzed in this study constitutes a valuable tool for assessing the performance of English language teachers and supporting continuous professional improvement. The results are consistent with recent international evidence recognizing systematic classroom observation and reflective feedback as essential strategies for enhancing instructional quality, promoting teacher professional learning, and improving educational outcomes, particularly within foreign language teaching contexts (Donaghue, 2026; Padgett, 2023; Reños & Pontillas, 2024; Thida & Yang, 2026; Acar et al., 2023).

CONCLUSIONS

The findings of this study indicate that classroom observation and teacher performance evaluation constitute essential mechanisms for promoting the continuous improvement of English language teaching, particularly when supported by instruments specifically designed to address the distinctive characteristics of this disciplinary field.

The study also demonstrated that specialized observation instruments enable a more relevant and accurate assessment of teaching practice by incorporating criteria associated with the pedagogical and communicative processes inherent in foreign language instruction.

Furthermore, the feedback generated through classroom observation emerged as a key component in fostering professional reflection, strengthening teacher development, and supporting decision-making processes aimed at improving educational practices.

Finally, the results highlight the importance of consolidating evaluation processes grounded in a formative approach, supported by evidence-based assessment and oriented toward professional learning. Such processes have the potential to enhance the quality of English language teaching and contribute to improved educational outcomes.

REFERENCES

- Acar, S. P., Akgün Özpolat, E., & Çomoğlu, İ. (2023). Teacher-tailored classroom observation for professional growth of EFL instructors: An exploratory case study. *Journal on Efficiency and Responsibility in Education and Science*, 16(1), 26-35. <https://doi.org/10.7160/eriesj.2023.160103>
- Alghamdi, N., Al-Nofaie, H., & Alkhamash, R. (2023). Peer observation of teaching in EFL classrooms in Saudi universities: Challenges and applications. *Arab World English Journal*, 14(2), 3-18. <https://doi.org/10.24093/awej/vol14no2.1>
- Amemasor, S. K., Oppong, S. O., Ben-Bright, B. G., & Essel, D. D. (2025). A systematic review on the impact of teacher professional development on digital instructional integration and teaching practices. *Frontiers in Education*, 10. <https://doi.org/10.3389/feduc.2025.154103>
- British Council. (2025). CPD framework and teacher observation. TeachingEnglish. <https://www.teachingenglish.org.uk>
- Brown, H. D. (2014). *Teaching by principles: An interactive approach to language pedagogy* (4th ed.). Pearson Education.
- Donaghue, H. (2026). Post observation feedback in language teaching: A review of recent research. *Language Teaching*. Cambridge University Press. <https://doi.org/10.1017/S0261444826101165>
- Inwang, J. (2024). Teacher learning through post-observation feedback in language education. En H. Donaghue (Ed.), *Post observation feedback in language teaching: A review of recent research*. *Language Teaching*.
- Kobayashi, Y., & Kobayashi, M. (2024). Research on teacher socialization and learning through post-observation feedback. En H. Donaghue (Ed.), *Post observation feedback in language teaching: A review of recent research*. *Language Teaching*.
- Lewis, J., & Wagner, E. (2023). Research on professional socialization through observation and feedback in language teaching. En H. Donaghue (Ed.), *Post observation feedback in language teaching: A review of recent research*. *Language Teaching*.

- Padgett, A. C. (2023). Teacher observations as professional development opportunities (Doctoral dissertation, University of South Carolina). University of South Carolina Scholar Commons.
- Reños, G., & Pontillas, P. (2024). Classroom observation and teachers' professional development activities: Basis for intervention plan. *American Journal of Arts and Human Science*, 3(3). <https://doi.org/10.54536/ajahs.v3i3.3077>
- Tapia, A., & Evison, J. (2024). Professional development and reflective dialogue in observation feedback processes. En H. Donaghue (Ed.), *Post observation feedback in language teaching: A review of recent research*. Language Teaching.
- Thida, S. M., & Yang, P. (2026). Classroom observation as reflective practice: A pre-service TESOL teacher's professional development. *Frontiers in Education*, 11, 1606820. <https://doi.org/10.3389/feduc.2026.1606820>
- UCLA Teaching and Learning Center. (2025). Classroom observation for instructors. University of California, Los Angeles. <https://teaching.ucla.edu/resources/teaching-guides/classroom-observation-for-instructors/>
- Yaşar, E., & Gürbüz, N. (2023). Observation feedback and teacher development in language education. En H. Donaghue (Ed.), *Post observation feedback in language teaching: A review of recent research*. Language Teaching.